

Additional Data File 1: List of all IRG gene family members and related genes described in Bekpen et al. 2005

Gene name	Genesymbol/ID	Synonyms	Genomic sequences /Accession no.	cDNA or EST sequence Accession numbers	Notes
Mouse					
<i>Irga1</i>	<i>Irga1</i> MGI:1795294 MGI:1653512		AC132320 AC102225	BI658674 (NMRI, 5'EST, nearly 100%) BG915086 (NMRI, 5'EST; not 100%)	
<i>Irga2</i>	<i>Irga2</i> MGI:915200 MGI:1257137 MGI:1257136		AC132320 AC102225 XM_140378	AA968296 (C57BL/6, 5'EST, 100%, not full length) AA968378 (C57BL/6, 3'EST, 100%, not full length)	Inducible by IFN-γ.
<i>Irga3</i>	<i>Irga3</i> New gene		AC132320 XM_140379 (C57BL/6J)	BY751179 (NOD, EST, not 100%, 610bp)	Inducible by IFN-γ.
<i>Irga4</i>	<i>Irga4</i> New gene		AC132320 XM_140380 (<i>Irgb4/Irgb5</i> tandem)	BY750970 (NOD, EST, nearly 100%, 700 bp) BU696309 (C57BL/6, EST, nearly 100%, 530 bp)	Inducible by IFN-γ.
<i>Irga5</i> Ψ	<i>Irga5</i> New gene		AC132320	None	A transcript is inducible by IFN-γ but the coding sequence of the gene is disrupted repeatedly.
<i>Irga6</i>	<i>Irga6</i> MGI:1926259 MGI:2147195 MGI:2147350	IIGP, IIGP1, Iigp1	AC135638	AJ007971 (C57BL/6, 100% correct) AF194871 (C57BL/6, also NM_021792, 100%) BC004649 (C57BL/6, cDNA 100%, 2330bp)	[13] (MGI:1889878); [16] Inducible by IFN-γ.
<i>Irga7</i>	<i>Irga7</i> New gene		NT_039674 (C57BL/6J, Chr.18 genomic contig, 73.9 Mb) XM_487533 (C57BL/6, 100%)	None known	
<i>Irga8</i>	<i>Irga8</i> MGI:953940 (C57BL/6) MGI:2384767 MGI:1489193 (CZECHII)	MGC:28198 BC023105	AC135638	BC023105 (CZECHII cDNA, = NM_145357, not 100%, full length) BB637466 (C57BL/6J, 5'EST, not 100%, not full length) BF163606 (CZECHII, not 100%, not full length) BE198503 (C57BL/6, 3'EST, 100%, not full length) BE198089 (C57BL/6, 3'EST, 100%, not full length) BX520309 (C57BL/6, 3'EST, 100%, not full length)	In C57BL/6 a non-canonical guanine after bp 849 in BC023105 (= aa 204) puts the sequence out of frame just before Helix H4; the reading frame is complete in BC023105 (CZECHII, <i>Mus musculus musculus</i>). Inducible by IFN-γ.
<i>Irgb1</i>	<i>Irgb1</i> MGI:1519766		AL645849	BC022776 tandem <i>Irgb2/Irgb1</i> (CZECHII, not 100%, protein: Q8R5D8) BF144722 (CZECHII, EST, not 100%, starts with 3' end of <i>Irgb2</i>)	The <i>Irgb2/Irgb1</i> gene pair is almost certainly transcribed in tandem. The protein has not yet been described. Inducible by IFN-γ.
<i>Irgb2</i>	<i>Irgb2</i>		AL645849	BC022776 tandem <i>Irgb2/Irgb1</i> (CZECHII, not	See note above, <i>Irgb1</i> .

	MGI:1518599			100%, protein: Q8R5D8) BF144934 (CZECHII, 5' <i>Irgb2</i> cDNA, not 100%) BY735436 (from cell line RCB-0527 Jyg-MC(B), strain unknown, 5' <i>Irgb2</i> , not 100%)	
<i>Irgb3</i>	<i>Irgb3</i> MGI:1553791 (FVB/N)		AL627237 AL669850 (unordered) AF060196 (129/SvJ, genomic, 1 bp difference, ATG(<i>Irgb3</i>)= bp 1353; Stop = bp 2659)	BF539106 (FVB/N, 3'EST, not 100%)	The genomic sequence of <i>Irgb3</i> is followed after 950 bp by a retroposon corresponding to the proteasome regulator PA28b [80] (MGI:1331589). The presence or absence of this retroposon unambiguously distinguishes <i>Irgb3</i> from <i>Irgb4</i> .
<i>Irgb4</i>	<i>Irgb4</i> MGI:1795392 MGI:3041173	9930111J21Rik	AL627237 AL669850 (unordered)	BC066104 (C57BL/6, <i>Irgb5/Irgb4</i> tandem, 100%) BI655221 (NMRI, EST, not 100%)	See note above for <i>Irgb3</i> . <i>Irgb4</i> is probably normally expressed as a distinct 3' exon in a tandem transcript downstream of <i>Irgb5</i> .
<i>Irgb5</i>	<i>Irgb5</i> MGI:3041173 MGI:2401562	9930111J21Rik	AL627237 AL645688 AL669850 (unordered)	BC066104 (C57BL/6, <i>Irgb5/Irgb4</i> tandem; not 100% at 5' end) AK037088 (C57BL/6, cDNA, = NM_173434, 100%, unknown 5' end) (protein = BAC29698= Q8CB10)	<i>Irgb5</i> is probably normally expressed as a separate 5' exon in a tandem transcript upstream of <i>Irgb4</i> . However AK037088 does not splice into <i>Irgb5</i> . Thus <i>Irgb5</i> can exist as a single p47 unit or as a tandem with <i>Irgb4</i> . The reference number MGI:2401562 refers to several ESTs belonging to <i>Irgb5</i> and <i>Irgb9</i> . Inducible by IFN-γ.
<i>Irgb6</i>	<i>Irgb6</i> MGI:98734 MGD-MRK-15077	TGTP, Mg21, Gtp2	AL627237 AL645688 AL669850 (unordered)	L38444 (C57BL/6, 100%) NM_011579 (NOD, 2 aa difference) U15636 (C.D2-Idh-1/Pep-3, 2 aa difference) BC085259 (NMRI, cDNA, 100%) BC034256 (CECHII, cDNA, not 100%)	[81, 82] Inducible by IFN-γ.
<i>Irgb7</i> Ψ	<i>Irgb7</i> New gene		AL645688 AL669850 (unordered)	None known	Pseudogene: STOP codon before G-domain. Not inducible by IFN-γ, no known transcript.
<i>Irgb8</i>	<i>Irgb8</i> MGI:1672892		AL645849	BG974191 (NMRI, 3' EST, not full length, not 100%,)	So similar to <i>Irgb1</i> , <i>b3</i> and <i>b4</i> that non-identical EST sequences are hard to disentangle.
<i>Irgb9</i>	<i>Irgb9</i> MGI:2401562 New gene		AL645849 XM_204704 (C57BL/6, full length, 100%)	BB630182 (EST, short)	The reference number MGI:2401562 refers to several ESTs belonging to <i>Irgb5</i> and <i>Irgb9</i> .
<i>Irgb10</i>	<i>Irgb10</i> MGI:1282384		AL928857	AI122314 (C57BL/6, short EST, not 100%)	Short, terminates before end of G domain in S6. Inducible by IFN-γ.
<i>Irgc</i>	<i>Irgc</i> New gene	CINEMA	AC073810 (RP23-57J6) GENSCAN00000140134	BB615720 (C57BL/6 cDNA, 99%, 606 bp) 36 ESTs, none full length (e.g. CA464745 5'mRNA, 874bp, 100% except of first two bp)	An <i>Irgc</i> -related sequence has recently been named HGTP-47 [4]. This sequence (NP_950178=NM_199013= AK089224, NOD) contains 4 frameshifts relative to the C57BL/6 genomic sequence leading to a largely incorrect protein sequence. The reference numbers MGI:2685948 and MGI:2685320 both relate to this

					error sequence.
<i>Irgd</i>	<i>Irgd</i> MGI:99448 MGD-MRK-16217	IRG-47, IRG47, Ifi47, 47kDa, Iigp4	AL645688 AL669850 (unordered)	M63630 (B6D2F1, =NM_008330, 100% correct)	[60]. This is the first report of a p47 GTPase and has given its name (IRG-47) to the whole family. Inducible by IFN- γ .
<i>Irgm1</i>	<i>Irgm1</i> MGI:107567 MGD-MRK-36139	LRG-47, LRG47, Ifi1, Iigp3	AL645849	U19119 (BALB/c, =NM_008326, 100% correct)	[83]; Two 5' splice variants exist. See notes human IRGM below. Inducible by IFN- γ .
<i>Irgm2</i>	<i>Irgm2</i> MGI:1926262 MGI:2144195	GTPI Iigp2	AL928857	AJ007972 (C57BL/6; 100%) NM_019440 (CZECHII, = BC005419, not 100%)	[13], MGI:1889878. Two 5' splice variants exist. Inducible by IFN- γ .
<i>Irgm3</i>	<i>Irgm3</i> MGI:107729 MGD-MRK-36305 MGI:2144580	IGTP Igtp	AL928857	U53219 (C57BL/6, cDNA, 100%) NM_018738 (NOD, cDNA, not 100%)	[84], MGI:82341 Inducible by IFN- γ .
<i>Irgq</i>	<i>Irgq</i> MGI:2667176	FKSG27	AC073810	AF322649 (C57BL/6, mRNA, = NM_153134)	
Human					
<i>IRGC</i>	UniGene Hs.515444 R30953_1 GeneID: 56269	CINEMA human IIGP5, cinema1	AC005622 HChr.19 cosmid	BC066939 (cDNA, 100%) NM_019612 (cDNA, 100%)	
<i>IRGM</i>	UniGene Hs.519680 GeneID: 345611 MIM: 608212	human LRG-47-like protein (LRG47, LRG-47), IFI1	AC010441 Chr.5 XM_293893 (splice variant a, 100%)	BC038360 (splice variant c, 3'EST) BC038539 (short EST) BI764111 (short EST) Sequences have been confirmed by RT-PCR (unpublished)	5 different 3' splice variants (a-e) (see main paper Bekpen <i>et al</i> , Fig. 8b). The orthology of <i>Irgm1</i> with human <i>IRGM</i> implied by use of the name <i>LRG47</i> or <i>IFI1</i> for the human gene is incorrect. The use of LRG47 as a synonym or alias for human IRGM is therefore not recommended.
<i>IRGQ</i>	UniGene Hs.546476 GeneID: 126298	Homo sapiens FKSG27, Irgql	AC006276	AF322648 (=NM_001007561 mRNA, 100%)	
Dog					
<i>IRGB11</i>	New gene		AACN010148430 AAEX1030324 AAEX1030325		
<i>IRGB12</i>	New gene		AACN01030937 AAEX1030324 AAEX1030325		Confirmed by RT-PCR but not sequenced. Inducible by IFN- γ .
<i>IRGC</i>	New gene	CINEMA	AACN010031536		

			AAEX01054272		
<i>IRGD</i>	New gene		AAEX01030325		
<i>IRGM4</i>	New gene		AAEX01059458		Confirmed by RT-PCR but not sequenced. Inducible by IFN- γ .
<i>IRGM5</i>	New gene		AACN010384735 AAEX01030325		Confirmed by RT-PCR but not sequenced. Inducible by IFN- γ .
<i>IRGM6</i>	New gene		AACN010300899 AAEX1030325		Confirmed by RT-PCR but not sequenced. Inducible by IFN- γ .
Fugu					
<i>irgf5</i>			Fugu_Sc2554 (Ensembl v3)		<i>irgf</i> genes of zebrafish, Fugu and Tetraodon have the long coding exon broken by an intron.
<i>irgf6</i>			Fugu_Sc2554 (Ensembl v3)	CA589084 (GI:25133662; 606 bp mRNA linear EST; hab53f04.y1 Fugu UT7 adult skin Takifugu rubripes cDNA clone) AL837863 (GI:21879801; 491 bp mRNA linear; F000A Takifugu rubripes cDNA clone F000A03aF7, mRNA sequence, skin)	See note above, <i>irgf5</i>
Tetraodon					
<i>irgf7</i>			SCAF112 (Ensembl v32, Jul 05)	GSTENT00000024001	<i>irgf</i> genes of zebrafish, Fugu and Tetraodon have the long coding exon broken by an intron.
<i>irgf8</i>			SCAF112 (Ensembl v32, Jul 05)	GSTENT00000023001	See note above, <i>irgf7</i> .
Zebrafish					
<i>irge1</i>	XP_693404		AL935330 (CH211-230C14) CR391937 (CH211-175G6)	BM316215 (3' EST)	Zebrafish <i>irge</i> genes have the long coding exon unbroken by an intron, like the mammalian p47 genes XP_693404 (GI:68383735, 502 aa linear VRT 30-JUN-2005 predicted: similar to immunity-related GTPase family, cinema 1 [Danio rerio]. DBSOURCE REFSEQ: accession XM_688312.1 (Short N-terminus)
<i>irge2</i>	XP_693474		AL935330 (CH211-230C14) CR391937 (CH211-175G6)	None	See note above, <i>Irge1</i> . XP_693474 (GI:68383738, 352 aa linear VRT 30-JUN-2005 predicted: similar to immunity-related GTPase family, cinema 1 [Danio rerio]. DBSOURCE REFSEQ: accession XM_688382.1 (Short N-terminus)
<i>irge3</i>			AL935330 (CH211-230C14) CR391937 (CH211-175G6)	AW233145 (5' cDNA)	See note above, <i>Irge1</i> .
<i>irge4</i>	XP_693622		AL935330 (CH211-230C14) CR391937 (CH211-175G6)	CN501017 (5' EST) CK142408 (5' EST)	See note above, <i>Irge1</i> . XP_693622 (GI:68383741, 385 aa linear VRT 30-

					JUN-2005 predicted: similar to immunity-related GTPase family, cinema 1 [Danio rerio]. DBSOURCE REFSEQ: accession XM_688530.1
<i>Irge5</i>	XM_681093		NW_635044 (GI:67045019; chr. 9 contig; bp 307225 308757)		See note above, <i>Irge1</i> . XM_681093 (GI:68365895, 1533 bp mRNA linear VRT 30-JUN-2005 predicted: Danio rerio similar to immunity-related GTPase family, cinema 1 (LOC557936), mRNA.
<i>Irge6</i>	XM_695163		NW_633868 (gi:67045754; chr. 18 contig; bp 5057602-5058696)		See note above, <i>Irge1</i> . XM_695163 (GI:68390584, 1095 bp mRNA linear VRT 30-JUN-2005 predicted: Danio rerio similar to immunity-related GTPase family, cinema 1 (LOC571560), mRNA.
<i>irgf1</i>	XP_700498		CR384077 DKEY-79I2	CN503005 (5' EST)	<i>irgf</i> genes of zebrafish, Fugu and Tetraodon have the long coding exon broken by an intron. XP_700498 (397 aa linear VRT 30-JUN-2005 predicted: similar to immunity-related GTPase family, cinema 1, partial [Danio rerio].
<i>irgf2</i>			CR384077 DKEY-79I2	None	See note above, <i>irgf1</i> .
<i>irgf3</i>			WGS traces zDH64-1061h13.q1k ZDH88-124d21.p1k zfish35935-195b06.p1c	AL924569	See note above, <i>irgf1</i> .
<i>irgf4</i>			ENSDARG00000010545	None	See note above, <i>irgf1</i> .
<i>irgg</i>			AL935330 (CH211-230C14) CR391937 (CH211-175G6)	CA473205 (5' EST)	No intron in long coding exon. Short, terminates in Helix F. Probably the 5' end of a tandem with <i>irgg1</i> .
<i>irgg1</i>			AL935330 (CH211-230C14) CR391937 (CH211-175G6)	BQ481364 (5' EST) and BQ481122 (3' EST) from cDNA clone IMAGE:5899497. The 5' end of this clone is in the 3' end of <i>irgg</i> and reads into the 5' end of <i>irgg1</i> .	Short, terminates in helix F. Probably the 3' end of a tandem with <i>irgg</i> .
<i>irgg2</i>	XP_684591		BX072550 DKEY-245P1	BF938149 5' EST and BI880124 3'EST from cDNA clone IMAGE:4200886	XP_684591 (GI:68381188, 379 aa linear VRT 30-JUN-2005 predicted: similar to RGD1311107_predicted protein [Danio rerio]. DBSOURCE REFSEQ: accession XM_679499.1
<i>irgg3</i>			BX127973; SP6 end of BAC DKEY-279M7 Zv4 scaffold1709.9	None	
<i>C. elegans</i>					
C46E1.3	WP:CE34758 GI:3300129;		AL008867.1 (GI:3217208, cosmid C46E1)	None	Predicted protein, tandem G domains.

	CAE17750				
W09C5.2	CAB63329.1 GI:6580259		Z82077 (GI:3873420, Cosmid W09C5)	None	Predicted protein.
Bacteria					
BAA10832	GI:1001345		BA000022.2 (GI:47118304, Synechocystis sp. PCC 6803)		Synechocystis sp . PCC 6803 Predicted protein.
BAA18140	GI:1653224		BA000022.2 (GI:47118304, Synechocystis sp. PCC 6803)		Synechocystis sp . PCC 6803 Predicted protein.
BAA18642	GI:1653731		BA000022.2 (GI:47118304, Synechocystis sp. PCC 6803)		Synechocystis sp . PCC 6803 Predicted protein.
BAC08557	GI:22294728		BA000039.2 (GI:47118315, T. elongatus BP-1)		Thermosynechococcus elongatus BP-1 Predicted protein.
BAC08842	GI:22295014		BA000039.2 (GI:47118315, T. elongatus BP-1)		Thermosynechococcus elongatus BP-1 Predicted protein.